

# **FAA Partie 1**

Honoré Nintunze

20/03/2016

## 0.1 TP1: Calcul de performance

Le but de ce TP est de calculer des mesures de performance pour une droite donnée  $y = 2x + 3$ .

Les résultats des mesures que j'ai obtenus sont:

$$J_{abs} = 0.73987984094$$

$$J_{l1} = 0.0896787983772$$

$$J_{l2} = 0.804228687838$$

$$J_{l\infty} = 2.51624302238$$

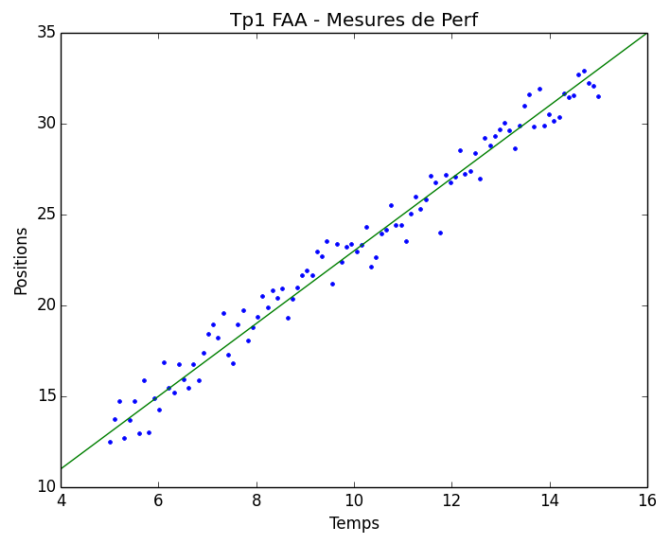


Figure 1: Droite donnée  $2 * x + 3$  et les points représentant les données

On constate que  $J_{l2}$  donne la meilleure mesure de performance.

## 0.2 TP2: Calcul des moindres carrés

Cette fois il ne s'agit plus de comparer les mesures de performance mais de trouver la droite qui approxime le mieux les données.

Pour cela, on cherche donc à minimiser l'erreur quadratique moyenne (MSE) par la méthode des moindres carrés, ce qui nous permet de trouver les paramètres  $\theta$  à partir des données.

Le calcul nous donne la droite  $y = 1.95293789x + 3.59623499$ .

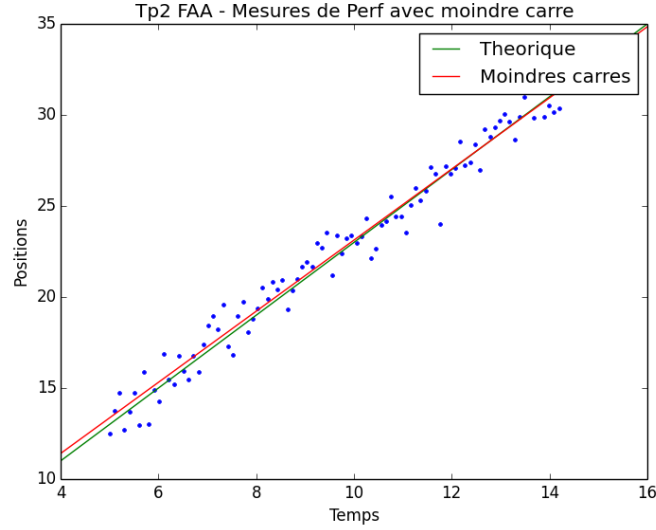


Figure 2: Ajustement des paramètres pas les moindres carrés

Les résultats des mesures que j'ai obtenus avec les nouveaux paramètres sont:

$$J_{abs} = 0.727356264922$$

$$J_{l1} = 0.0877279862436$$

$$J_{l2} = 0.769619957035$$

$$J_{l\infty} = 2.55866627298$$

La différence entre les résultats du TP1 et ceux des moindres carrés donne:

$$diff(J_{abs} - J_{abs}MC) = 0.0125235760181$$

$$diff(J_{l1} - J_{l1}MC) = 0.00195081213361$$

$$diff(J_{l2} - J_{l2}MC) = 0.0346087308024$$

$$diff(J_{l\infty}MC - J_{l\infty}) = 0.0424232506014$$

### 0.3 TP3: Descente de gradient

On peut faire une descente de gradient pour essayer de trouver un minimum local (en espérant trouver le minimum global) pour trouver les paramètre qui approximent au mieux les données.

J'ai donc implémenté une descente de gradient globale et une descente de gradient stochastique.

La descente de gradient globale approxime la droite  $y = 1.95461147x + 3.57808937$  pour une mesure normale 2 de 0.76964575673 alors que la descente de gradient stochastique donne  $y = 2.19790012x + 0.34082447$  avec le tirage que j'ai effectué, pour une mesure normale 2 de 1.92907081566.

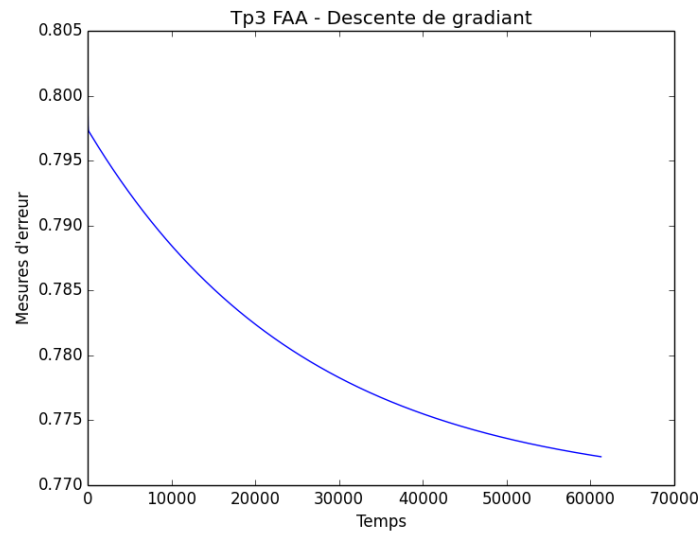


Figure 3: Courbe de la descente de gradient globale

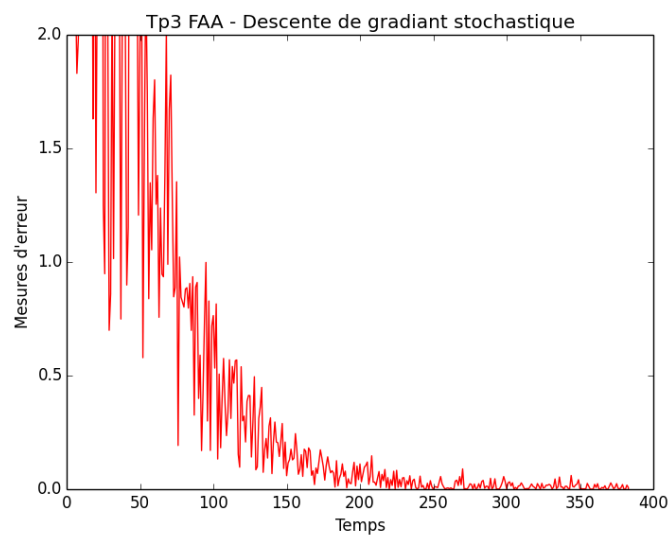


Figure 4: Courbe de la descente de gradient stochastique

En faisant varier le paramètre  $\alpha$  , on peut observer les courbes suivantes:

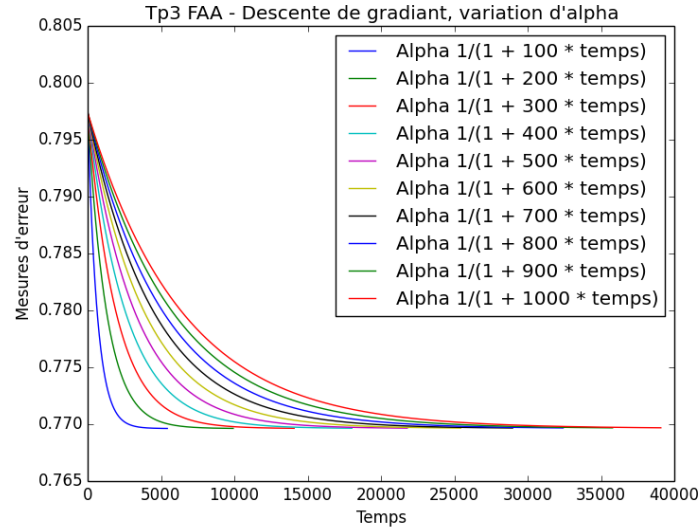


Figure 5: Variation du paramètre  $\alpha$  pour la descente globale

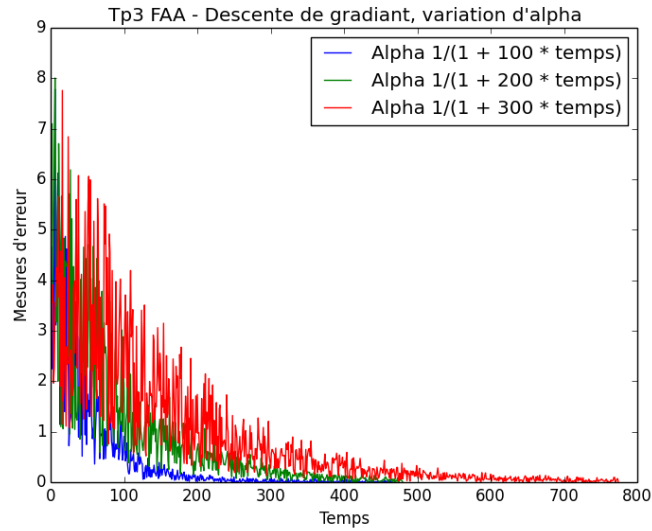


Figure 6: Variation du paramètre  $\alpha$  pour la descente stochastique

## 0.4 TP4: Vision statistique de l'apprentissage

Pour ce TP, le but est d'approximer des familles de fonction pas évidentes pour lesquelles il faut enrichir l'espace d'hypothèse.

On avait donc 3 jeux de données et j'ai pour les 3 jeux approximer des fonctions en complexifiant l'espace par des polynômes de degré de plus en plus grand.

Pour trouver l'approximation, on utilise les moindres carrés:

- Pour le jeu 1:  $y = 2.01633307x + 2.94585478$
- Pour le jeu 2:  $y = 1.00024654x^2 + -0.99260885x - 0.25839725$
- Pour le jeu 3:  $y = -15.97924772x^{10} + 34.70984613x^9 + 119.64487763x^8 - 194.43573825x^7 - 370.19318234x^6 + 494.51831655x^5 + 514.56263705x^4 - 512.87538799x^3 - 261.34180207x^2 + 144.47932936x^1 + 21.31140909$

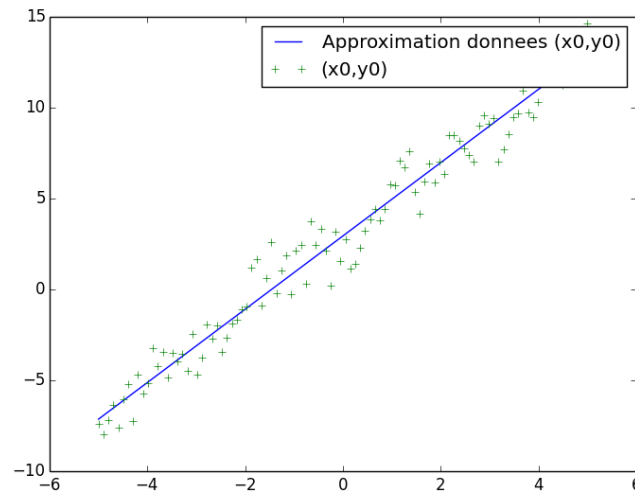


Figure 7: Jeu de donnée 1

Je n'ai pas réussi à estimer le risque réel par validation croisée à K plis.

## 0.5 TP5: Classification par régression logistique

Pour ce TP il fallait classer des hommes et des femmes par leur taille et leur poids en apprenant la relation derrière le jeu de donnée, ceci par régression logistique.

Ici, tout ce que j'ai pu faire c'est afficher les 2 classes (hommes et femmes) sur une figure. Je n'ai pas réussi à afficher des séparateurs linéaires corrects, aucuns ne coupent les données. je pense que ma descente de gradient avec sigmoïde ne marche pas mais je ne sais pas pourquoi. Je n'ai donc pas fait la régression sur 2 dimensions.

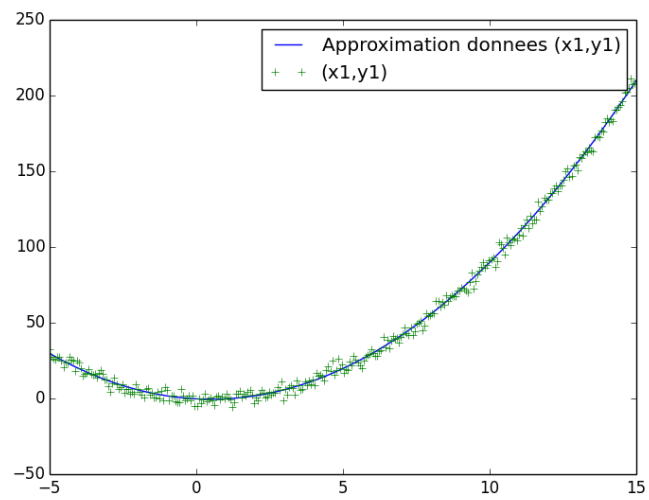


Figure 8: Jeu de donnée 2

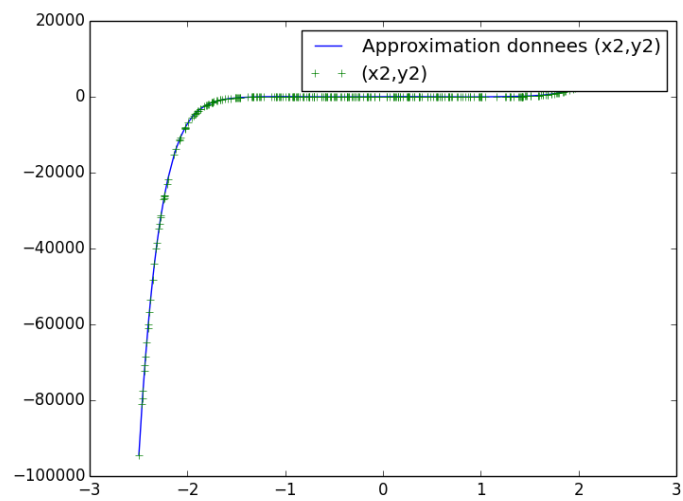


Figure 9: Jeu de donnée 3

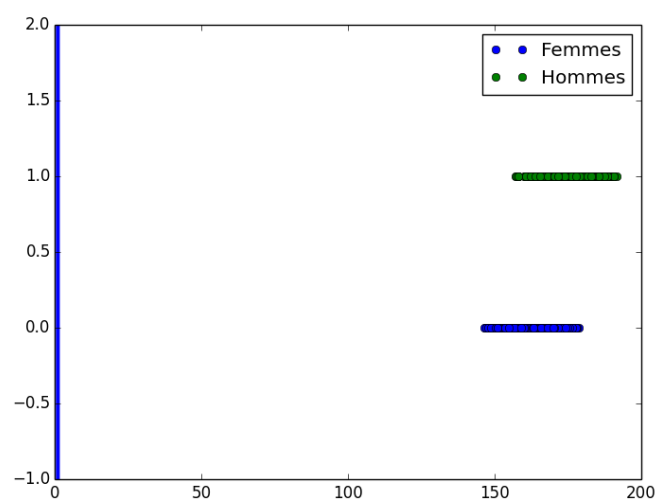


Figure 10: Données des classe H/F avec séparateurs linéaires incorrects